Restoration of deleted and assimilated consonant sequences in conversational French speech: effects of preceding and following context

Danielle Duez
Laboratoire Parole et Langage, CNRS UMR 6057,
Université de Provence
duez@lpl.univ-aix.fr

The present study was designed to test the effects of preceding and following context on the restoration of deleted and assimilated consonant sequences in conversational French speech. Consonant restoration was found to be enhanced by lexical information and by the different sources of information contained in the breath group. The fundamental contribution of breath-group information suggests that the breath group, which is a production unit in conversational speech, might be a perception unit. Interestingly, the results also appear to be in line with Lindblom's claim (1990, 1996) that speakers adapt their utterances to the listeners' needs and maintain sufficient contrast for promoting correct lexical access.

1 Introduction

In conversational speech, which is considered a good example of hypospeech, segments are often reduced, altered, omitted or combined with other segments, compared to segments in the same words carefully produced. However, in spite of being hypo-articulated, such speech segments may still be sufficiently rich for perception: speakers are guided in their speech by their assumptions about their listeners' experience and linguistic knowledge so that they can adapt their utterances to the listeners' needs and maintain sufficient contrast for promoting correct lexical access (Lindblom 1990, 1996). In turn, listeners bring their knowledge and their linguistic experience to bear in order to correctly identify utterances and understand speech.

Evidence that listeners identify speech sounds in part by normalizing them with respect to their phonetic context has been found in a large number of studies. For example, Ohala et al. (1981) (summarized in Ohala 1986) found that in steady-state (transitionless) vowels from a synthetic /i/ and /u/ continuum, the crossover between these two vowels as judged by American English listeners was closer to the front when embedded in the context of alveolar consonants (/s-t/) than when flanked by labial consonants (/f-p/). Similarly, Beddor et al. (1986) showed that ambiguous vowel items usually identified as /ɛ/ were more often identified as /æ/ when nasalization was added, but that this perceptual height shift did not occur when the nasalized vowels were in a nasal consonant context. Compensatory effects were also demonstrated by Mann & Repp (1980), who reported a shift in the /s-ʃ/ boundary as a

DOI:10.1017/S0025100301001086 Journal of the International Phonetic Association (2001) 31/1

function of vowel context: fricatives were more often labeled as /s/ when followed by [u] than by [a].

An even stronger role of context was shown in studies dealing with the restoration of a missing phoneme, also called 'phonemic restoration' (Warren 1970). Listeners typically reported as intact an utterance partly replaced by a cough; they appeared to perceptually restore the missing phoneme. Later, phonemic restoration was also shown to depend on the spectral composition of the replacement sound (Warren & Obusek 1971) and on the acoustic similarity between the missing sound and the replacing one (Samuel 1981, 1987). For example, when white noise was the replacement sound, fricatives were better restored than vowels, whereas the opposite pattern was found with pure tone replacement (Samuel 1981).

The role of adjacent segments in phonemic restoration has been demonstrated in studies using a splicing method. For example, Ali et al. (1971) showed that subjects could predict the presence of nasal consonants in (C)CVC syllables from which the final consonant (nasal or oral) and its vocalic transition had been spliced out. Using the same method, Benguerel & Adelman (1975) demonstrated that a missing final [i], [y] or [u] of a VCCCV sequence was identified well above chance level. Information on a speech segment is not restricted to the speech segment itself but is also retained in adjacent speech segments. In turn, the acoustic cues of contextual segments are used in phoneme identification and may help restore missing elements.

Lexical knowledge also exerts a strong influence on phonemic restoration. Samuel (1981, 1987) observed more restoration for high-frequency words than for low-frequency words, for long words than for short ones and for word-final sounds than for word-initial ones. These results support the view that access to more frequent words is more efficient (Rubenstein & Pollack 1963, Foss 1969) and that lexical access is dominated by word beginnings (Cole 1973).

Restoration of missing phonemes may not require the construction of semantic, syntactic or other higher-level redundancies (Ohala & Feder 1994). Listeners' normalization of vowel quality (for vowels in the /i/-/u/ continuum) as a function of the consonant that followed (/d/ or /b/) when either physically present or replaced by noise was found to be affected by the perceived consonant, even if it was replaced by noise. The linguistic categorization of the adjacent consonant guides listeners' judgements: they are able to integrate linguistic, i.e. categorical, information into recognition.

The findings on phonemic restoration clearly demonstrate that processes independent of the acoustic signal, namely linguistic and other factors, play a crucial role in the perception and comprehension of speech. This is of interest for conversational speech perception, since conversations are often produced in noisy environments where speech segments are altered or deleted. It may be assumed that listeners use contextual information and linguistic knowledge to restore assimilated, distorted or omitted speech sounds.

This assumption was tested in a perceptual investigation of consonant sequences that contained a deleted consonant or a consonant changed into another consonant (Duez 1998). The following two questions were addressed: 1) Are deleted and assimilated consonants restored when presented in context? 2) What level(s) is (are) involved in consonant restoration, i.e., does the immediate acoustic-phonetic context allow listeners to compensate for reduced information, or does restoration operate at a higher level, such as the syllable, word or phrase? The successful restoration of assimilated and reduced consonants in contexts would confirm Lindblom's claim (1990) that phonetic gestures are not made more distinctive than they need to be, and that speakers produce phonetic contrasts that are sufficient for lexical access.

Without taking any position on the actual processes involved, in this paper we use the term 'restoration' as a cover term for processes which, at any linguistic level, allow

the listener to identify the underlying/phonological form from an impoverished realization.

We used the gating technique (Pollack & Pickett 1963; Grosjean 1980, 1985) to present consonant sequences in context and to determine what quantity of contextual information is necessary for the restoration of the underlying/phonological consonant sequence. In the gating paradigm, words are presented to subjects in increments, which gives us good control over the temporal span of the acoustic-phonetic information listeners hear before responding. In the present study, context effects where limited to breath groups containing consonant sequences. Vaissière (1991) defined a breath group as a prosodic unit ending with significant syllable lengthening and an *F0*-rise or *F0*-fall contour. Most often there is also presence of a silent pause. Breath groups mostly correspond to syntactic phrases and sense groups (as defined by Grammont 1914). Breath-group choice was intended to (at least partly) neutralize variability in gated-interval length. Listeners were provided with the left context only (the context preceding the consonant sequence back to phrase beginning) or with both left and right contexts. Left and left+right contexts were gated syllable by syllable, except for the first two gated intervals, which consisted of the complete preceding and/or following vowel, and the syllable to which the reduced or assimilated consonant(s) belonged. Listeners' restoration of phonological consonant sequences was analyzed as a function of the quantity of linguistic information contained in the gated intervals.

2 Consonant-sequence characteristics

2.1 Perceptual criteria

For this study we selected forty-one consonant sequences (37 for the experiment and 4 for a pre-test) where the realized form was clearly different from the underlying/phonological sequence. Their selection was based on the results obtained in an investigation of hypoarticulation effects on two-consonant sequences (Duez 1998). The term 'two-consonant sequences' refers to sequences located in the phonological transcription of the conversational speech produced by two male speakers. There were consonants /C1C2/ occurring initially or finally in a syllable as in *style* and consonants occurring in two consecutive syllables within and across words. The latter consisted of a coda consonant plus an onset consonant /C1#C2/, as in *pulsations* (beats), or of two syllable-initial consonants separated by an optional schwa /C1(ə)C1/ as in *dans l*(*e*) *monde* 'in the world'. The non-realization of the optional schwa was established auditorily and acoustically. An utterance list is given in the appendix.

A total of 720 consonant sequences (360 per speaker) were extracted with their adjacent vowels and presented to sixteen naive listeners for identification. The listeners were told that the stimuli were excerpts from meaningful sentences and asked to write down everything they heard (clusters, single consonants, no consonant, vowels) using orthographic transcription. Then, to investigate the patterns within the identification responses each consonant sequence was assigned a single identification value corresponding to the phonological sequence reported by at least 75% of listeners (i.e. at least 12). This yielded two main groups: 1) sequences identified as the underlying/phonological sequences and 2) sequences identified as different from the underlying/phonological sequences. In the latter group, there were four cases: 2a) the phonological C1 and/or C2 was identified as another consonant (e.g. /f/ identified as /v/), 2b) one or two features were identified (e.g. /f/ identified as labial and/or fricative), 2c) neither a consonant nor a feature was identified, so that the consonant was reported as unidentified, and 2d) the consonant was omitted. To focus on the restoration process,

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

Table 1 Number of consonants in the test as a function of consonant type and pattern of assimilation or reduction.

	Stops			Fricatives				Sonorants			Total
	p	t	d	v	s	z	ʒ	n	ʀ	l	
Nasalized	-	2	5	–	–	–	–	–	–		7
Place-Assimilated	–	–	–	–	–	–	–	1	–		1
Devoiced	–	–	1	1	–	1	3	–	–	–	6
Voiced	1	–			3			–	–	–	4
Deleted	–	4	2	4	–	1			3	5	19
Total	1	6	8	5	3	2	3	1	3	5	37

the selected consonant sequences had to contain an omitted consonant or a consonant perceived as another consonant.

2.2 Acoustic criteria

Wide-band spectrograms were generated for each sequence. Some acoustic cues proven useful in the identification of sonorants and voiced and voiceless fricatives and stops were inspected visually in order to relate the perceptual data to the spectrographic data and to determine how accurately listeners interpret the acoustic cues. The results obtained for the 37 sequences used in the experiment are presented in table 1.

Nasalized consonants. Nasal consonants are characterized by mid-frequency formants (Fujimura 1962). The five /d/s reported as nasals had three visible formants. Their average values were 1170 Hz, 2000 Hz and 4000 Hz, and their duration ranged from 45 ms to 77 ms. There were two /t/s. One had a visible 3000-Hz formant and was 55 ms long. The other had two visible formants (1500 Hz and 2000 Hz) and its duration was 57 ms.

Voiced and devoiced consonants. Voiced fricatives and stops have low frequencies, unlike voiceless fricatives and stops. Voiced fricatives and stops are significantly shorter than their voiceless counterparts, especially voiced stops, which exhibit a shorter VOT (when present) than voiceless stops (VOT is the interval between burst onset and the first vowel period). The presence of these characteristics was checked for the different voiceless consonants reported as voiced, and vice versa. There was complete devoicing for the three /ʒ/s reported as /ʃ/, the /z/ reported as /s/, and the /v/ reported as /f/: two of the /ʒ/s were 50 ms long, the third one was 60 ms long, and the /z/ and the /v/ were 113 ms and 61 ms, respectively. Correspondingly, for the three /s/s reported as /z/ there was presence of a voicing bar. Their durations were 33 ms, 39 ms and 65 ms, respectively. Similar findings were obtained for stops: a /p/ identified as a /b/ exhibited low periodic frequencies and was 61 ms long ; there was a silence for a /d/ identified as /t/, with a duration of 63 ms (occlusion: 57 ms, VOT: 6 ms).

Deleted consonants. There were six omitted stops: 4 /t/s (3 C1s and 1 C2) and two /d/s (1 C1 and 1 C2). For the /d/ preceding a /z/ and the /t/ following an /s/ there was no trace of an occlusion or a burst. There was no burst for the /d/ followed by a /b/, the two /t/s followed by a /d/, and the /t/ followed by a /p/. However, the longer duration of the remaining /b/ (103 ms), /d/s (94 ms and 122 ms), and /p/ (114 ms) compared to the

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

Table 2 Number of assimilated and deleted consonants as a function of consonant-sequence type (homosyllabic and heterosyllabic), location within a sequence (C1 or C2 in boldface italics), and word class. [C1#C2]'s consist of a coda consonant plus an onset consonant, [C1#C1]'s of two syllable-initial consonants.

Class	Type	Heterosyllabic			Homosyllabic	
		C1#C2	***C1***#C1	C1#***C1***	***C1***C2	C1***C2***
Content words	Adjective	6	–	–	–	–
	Verb	4	–	2	1	–
	Noun	6	–	–	1	1
	Adverb	1	–	–	–	–
Function words	Preposition	1	3	1	–	–
	Pronoun	2	4	–	–	–
	Article	2	1	1	–	–
Total		22	8	4	2	1

mean reference duration of intervocalic /b/s (M = 66 ms, SD: 22 ms), /d/s (M = 61 ms; SD: 24 ms) and /p/ (M: 90 ms, SD: 13 ms) makes it seem likely that the consonants were subject to assimilation and not deletion. However, as they were reported as deleted they were considered as such. The duration of the remaining /s/ (M: 145 ms) compared to the reference /s/ (M: 104 ms, SD: 19 ms) in the sequence /st/ suggests the same situation. Five voiced fricatives were omitted (4 /v/s and 1 /z/): there was no trace of them when followed by /s/ or /f/ (2 /v/s and 1 /z/) or an /r/ (1 /v/ and 1 /(ʒ/). The remaining consonants were roughly of the same duration as the same intervocalic reference consonants. However, the remaining /f/ preceded by a /v/ was 145 ms while the reference mean of a single /v/ was 88 ms, which again suggests an assimilation process. There was omission of three /ʀ/s and five /l/s, none of which exhibited vocalic formants. One /n/ was reported as /m/ before a /p/. The four consonants on the pre-test were a /d/ reported as an /n/, a /d/ reported as a /t/, a /v/ reported as an /f/ and an omitted /t/ in an /st/ sequence. Their patterns were similar to those reported for the test consonants.

2.3 Word class and consonant location within syllables and phrases

Table 2 shows that the sequences were mainly heterosyllabic (34 out of 37). They consisted of a coda consonant plus an onset consonant [C1#C2], or of two syllable-initial consonants [C1#C1]. As already mentioned, the latter type resulted from the non-realization of an optional /ə/. In function words, there were more heterosyllabic sequences resulting from the absence of an optional /ə/ than sequences consisting of a coda consonant plus an onset consonant. In content words, it was the other way around. The three homosyllabic sequences belonged to content words. Roughly, the frequency of consonant sequences as a function of type, word class and location within the syllables reflects the fact that consonants are more sensitive to reduction and contextual assimilation in heterosyllabic sequences than in homosyllabic ones, in coda position than in onset position (Byrd 1996, Duez 1998). Note also that all deleted and assimilated consonants were located within phrases totally void of interruptions, filled pauses, repeats, false starts and drawls.

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

3 Method

3.1 Labeling

Sentences containing the selected consonant sequences were digitized at a sampling rate of 32 kHz with a computer. Wide-band spectrograms and oscillograms were generated for each sentence and labeling was performed. In the two experimental conditions (left-only context, left+right context), the first gated interval was always a consonant sequence with half of the preceding and following vowels (V1C1C2V2): the middle of V1 and V2 was located on oscillograms and spectrograms. In the next gated intervals, there was a complete preceding and/or following vowel: V1 onset was the first *F*1-period and V2 offset was the last *F*1 period. Other gated intervals were increased syllable by syllable. Syllable onset was defined as the beginning of voiced or voiceless noise for syllables beginning with a voiced or voiceless fricative, and as the closure-interval onset (voiced or silent) for voiced or voiceless stops, or the second formant for vocalic consonants. For V or VC syllables, it was the first period of the vowel. Syllable offset was defined as the point where noise disappeared for final fricatives, the release of the burst (if any) or the end of the closure interval for final stops, the point of maximal spectral change increase for vocalic consonants and the last period of phonation for vowels.

3.2 Stimuli

Two series of stimuli were prepared: a left-only context series and a series including both left and right contexts. In the left-only context series, all stimuli ended with the middle of the vowel (V2) following the C1C2 sequence. The first stimulus started with the middle of V1 (the preceding vowel), the second with V1 onset, and the third with the onset of the syllable containing V1. The subsequent stimuli were incremented syllable by syllable. In the left+right series, the first stimulus was again the V1C1C2V2 segment, the limits of the second stimulus were V1 onset and V2 offset, and the limits of the third stimulus were the onset of the syllable containing V1 and the offset of the syllable containing V2. The subsequent stimuli were incremented syllable by syllable until the whole breath group had been presented. There were 41 sets per series: four for the pre-test and 37 for the test. The number of stimuli ranged from two to nine. A 20-ms ramp was applied at the beginning and the end of the signal. The maximal attenuation at the end points was 20 dB. Each stimulus was repeated three times, and the interval between repetitions was 1 s. The interstimulus interval was three seconds plus the stimulus duration multiplied by two. The interval between sets was 7 s. Each series started with the four pre-test sets, and then the 37 test sets were presented randomly. Within a set, the stimuli were presented in an increasing order of duration, that is with increased left information and increased left plus right information. Each series was recorded on a CD and answer sheets were compiled for each series.

3.3 Subjects

Twenty French listeners with no hearing problems participated. They were students at the University of Provence and were paid for their participation. They were divided into two groups, one per series. Each subject participated in one series only. The experiment took place in a quiet room. The subjects were tested individually and listened through Sennheiser HD222 headphones. They were told that the stimuli were excerpts from conversational speech and were asked to write down everything they heard on their answer sheet using orthographic transcription. They were not allowed to go back on and change their answer after more context was provided.

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

3.4 Organization of the data

To determine the patterns within the restoration responses, each C1C2 sequence was assigned a single restoration value corresponding to the C1C2 sequence reported by at least 75% of the listeners (i.e. at least 8). A consonant was also considered to be partially restored when only one or two of the missing cues were restored. This was the case, for example, for a deleted /d/ restored as /l/ by the majority: the voicing and place features were restored but not the manner feature. Restoration responses were examined as a function of consonant type, the class of the word to which it belongs, the location of a target consonant within a sequence and a word (C1 or C2, final or initial), and the phonetic and linguistic information contained within the gated interval. The different levels of information were defined as follows:

1) Adjacent vowel level. This level was examined to determine the extent to which listeners relied on adjacent vowels to compensate for the loss of information in the reduced or assimilated consonant. Another objective was to test the validity of the results obtained in the perceptual analysis of consonant sequences. To focus on acoustic consonant cues, VCCV utterances were extracted from the speech signal and the vowels truncated. This incorrect vowel duration may have been crucially misleading for judging consonantal contrasts.
2) Syllable-level. Perceptual experiments (Miller 1986) have revealed that listeners adjust some consonant contrasts to syllable duration and changes in articulation rate in order to distinguish phonetic categories.
3) Word level. It has been assumed that lexical representation plays a role in 'undoing' place assimilation, since listeners have no problems correctly parsing and identifying words which are changed into non-words under assimilation effects (Lahiri 1995). Note that monosyllabic words such as *grand* and *crise* were considered as words, not as syllables.
4) Prosodic-word level. Prosodic words are widely accepted as units of production (Vaissière 1991, Wheeldon & Lahiri 1997), they are bounded on the right by syllable lengthening and mostly consist of an article plus a noun, or a personal pronoun plus a verb.
5) Breath-group level. The breath group is a major unit of conversational speech (Vaissière 1991). Moreover, most assimilated consonants have been found within breath groups (Duez 1995, 1998).

4 Results

4.1 Left-only context

There was no significant effect of word class [$F(1,124) = 0.6$, $p = 0.4$] but a significant effect of information level [$F(6,124) = 9.7$, $p = 0.000$] on consonant restoration. As can be seen in table 3, performance was rather low since 25 out of the 37 assimilated or deleted consonants remained unrestored. Partial restoration was observed for one deleted consonant: listeners restored a phonological /d/ as /l/ or /d/ at the point when they heard the complete syllable. At the prosodic-word level, there was also partial restoration for the final /t/ (assimilated as an /n/) of the word *différentes* (/difeʁɑ̃t/) in the phrase *de différentes couleurs* 'with different colors'. The two proposed consonants /d/ and /t/ could form a legal French word *différant de* or *différentes* 'differing from' or 'different'.

Five assimilated consonants and one omitted consonant were restored at the word level. As soon as listeners heard the complete word, they realized that the word-final consonant was not /f/, /n/, /b/, /z/ or /m/ and proposed /v/, /d/, /p/, /s/ and /n/, i.e. the

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

Table 3 Number of restored consonants as a function of word class and location within the syllable (boldface italics). The information levels were defined as follows: syllable (Syll), monosyllabic word (MSW), polysyllabic word (PSW) and prosodic word (PW). The results can be compared with the total number (TN) of assimilated and reduced consonants. No R stands for no restoration. The cases of partial restoration are given in parentheses.

Word Class	Location in Syllables	Number		Information Level			
		TN	No R	Syll	MSW	PSW	P W
Content	***C1***#C2	17	8	0	3	2	4(1P)
	C1C2	2	2		–	–	–
	C1***C2***	1	1		–	–	–
	C1#C1	2	2		–	–	–
Function	***C1***#C2	5	4	1	1	–	–
	C1#C1	8	6	2(1P)	–	–	–
	C1#***C1***	2	2		–	–	–
Total		37	25	2	4	2	4(1P)

consonant corresponding to the underlying phonological one. The example in (1) is a good illustration of the restoration of an assimilated consonant.

(1) *ils ne doivent pas* 'they must not'
Perceived: [ilnədwa̬pa] Restored: /ilnədwavpa/

The omitted consonant restored at the word level was the final /ʁ/ of the adverb *toujours*, illustrated in (2).

(2) *c'est toujours systématique* 'it's always systematic'
Perceived: [sɛtuʒuØsistematik] Restored: /sɛtuʒuʁsistematik/

The restoration of three other consonants occurred at the prosodic-word level. The information contained in the article was crucial in word disambiguation. In example (3), the vowel /i/ in the word *crise* 'crisis' had a long duration (106 ms), reflecting some lengthening effect of the following underlying /z/. However, in spite of the long vowel duration, which is known to be a voicing cue in some consonantal contrasts, the form [kʀiz̥] was compatible with the third person *crissent* (/kʀis/) of the verb *crisser* 'to crunch' or 'to screech' for the following reasons: 1) the absence of low frequencies, which is one of the primary cues for unvoiced fricative identification, and 2) the fact that the form [kʀiz̥] was presented out of context, i.e. without any possible reference to overall sentence duration. When [kʀiz̥] was preceded by the article *la* 'the', there was no longer any ambiguity and the listeners proposed the right word.

(3) *la crise fait* 'the crisis makes'
Perceived: [lakʀiz̥fe] Restored: /lakʀizfe/

There were 25 assimilated or deleted consonants that were not restored, three of which were C2s. There may be different reasons for this failure. The fact that the left-only context was limited to the beginning of the breath group containing the consonant sequence may be one of the reasons. As is shown in (4), the context of the previous sentence or phrase may also play a crucial role in consonant restoration. There is a

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

vowel-quality difference in the two forms of the verb *pouvoir* 'can'. The plural form has an open vowel [pœv] while the singular form has a closed vowel [pø]. There is also a strong tendency to neutralize the mid-vowel contrast /œ/-/ø/ in non-prominent syllables (Wioland 1991). In example (4), the word *peuvent* was produced with a vowel closer than usual, therefore, the answer proposed by the listeners for *qui peut* 'who can' (singular form) was compatible with the acoustic-phonetic information contained in the gated interval. If provided with the preceding sentence, the listeners would probably have proposed the verb's plural form *peuvent*.

(4) *qui peuvent faire ça* 'who can do that'
(preceding context: *Il y a des cireurs de chaussures professionnels* 'There are professional shoe-shiners'
Perceived: [kipœØfɛʁsa] Restored: /kipœvfɛʁsa/

The need for both left and right contexts in the restoration of missing cues may be another reason. In (5), listeners had to wait for the noun *décorations* to be sure whether the adjective's gender was masculine *hauts* or feminine *hautes*.

(5) *les plus hautes décorations* 'the highest decorations'
Perceived: [leplyoØdekoʁasjɔ̃] Restored: /leplyotdekoʁasjɔ̃/

A third reason may be that the acoustic-phonetic information contained in the utterances was too poor to allow for successful restoration.

4.2 Left+right context

Again, an ANOVA performed on listener responses did not reveal a significant effect of word class [$F(1,166)$, $p = 0.59$] but it did reveal a significant effect of information level [$F(5, 166) = 34$, $p = 0.000$]. Restoration rate was found to strongly depend on right context. As seen in table 4, 29 of the 37 consonants were completely restored when they had the left and the right context.

In general, the results obtained in the present section confirmed those obtained for

Table 4 Number of restored consonants as a function of word class, and location within syllable (boldface italics). The information levels were the syllable (Syll), monosyllabic word (MSW), prosodic word (PW), PBG (partial breath group), and breath group (BG). The results can be compared with the total number (TN) of assimilated and reduced consonants. No R stands for no restoration. The cases of partial restoration are given in parentheses.

Words	Location in syllables	Number	Information level					
		TN	No R	Syll	MSW	PW	PBG	BG
Content	***C1***#C2	17	2	0	1	5	2	7
	C1C2	2	0	–	–	–	1	1
	C1***C2***	1	1	–	–	–	–	–
	C1#C1	2	1	–	–	–	–	1
Function	***C1***#C2	5	2	–	–	1	2	–
	C1#C1	8	1	2(1P)	–	–	1	4
	C1#***C1***	2	1	–	–	–	–	1
Total		37	8	2	1	6	6	14

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

left-only context. Again, the adjacent vowel was not found to play a role in consonant restoration. The consonants previously restored with the left-only context were restored at the same information level with the left+right context. Two word-final consonants were restored in a gated interval consisting of an incomplete breath group. For example, the final /z/ of the numeral adjective *quinze* 'fifteen' was restored when the first two syllables of the following word *centimètres* 'centimetres' were presented to listeners. There was restoration of 14 consonants at the breath-group level. This is illustrated in (6) and (7).

(6) *C'était quelque chose d'inoui* 'It was something unheard of'
Assimilated: [setɛkɛlkəʃozØinui] Restored: /setɛkɛlkəʃozdinui/
(7) *On n'incite pas les gens* 'People are not encouraged'
Assimilated: [ɔ̃nɛ̃siØpaleʒɑ̃] Restored: /ɔ̃nɛ̃sitpaleʒɑ̃/

Eight consonants were not restored. As mentioned above, in some cases the failure may have been due to insufficient acoustic-phonetic information. This may be particularly true for the three cases where listeners were presented with the second version of a phrase, which is often shorter and reduced. In other cases, critical previous and/or subsequent information may have been missing. For example, as already shown in (4), listeners should have known with preceding context that the verb's subject was plural, not singular. The rarity of a word that is particularly hard to guess, such as *gent* may be a reason why its final consonant was not restored, as in (8). Finally, the reduced form of function words such as *il* ⇒ [i] and *je* ⇒ [ʃ] may be one of the phonological forms (9).

(8) *Ah la gent féminine* 'Ah the fair sex'
Assimilated: [laʒɑ̃nfeminin] Phonological: /laʒɑ̃tfeminin/
(9) *C'est qu'ils ont essayé* 'Well they tried'
Assimilated: [sɛkizɔ̃eseje] Phonological: /sɛkizɔ̃eseje/ or /sɛkilzɔ̃eseje/

5 Concluding remarks

The present study was designed with one basic question in mind: How critical a role does a listener's linguistic knowledge play in the restoration of assimilated and omitted consonants in conversational speech? The present results provide some answers to this question. About 80% of the C1C2 sequences reported as omitted or different from their phonological C1C2 counterparts when flanked with half the preceding and following vowels were totally restored when presented in breath groups. When all other sources of information were removed, listeners could only rely on the acoustic cues retained in the consonants. When listeners could use lexical, syntactic, prosodic and semantic information, they restored assimilated and omitted consonants. These results are in line with previous work on the perception of mispronunciations (Cole 1973, Cole & Jakimik 1978) and phoneme restoration (Samuel 1981, 1987). They provide a good illustration of so-called 'top-down processing': when the speech signal is noisy, ambiguous or incomplete, the receiver uses internal linguistic knowledge and an algorithm to guess and reconstruct the intended message (Ohala 1986).

The data from this experiment indicated high variability in the amount of information needed to restore assimilated and omitted consonants. There was little restoration in the left-only context condition. A few consonants were restored at the word level when listeners were presented with entire word and there were no other possible candidates. Others were restored at the prosodic-word level when the information contained in the prosodic words removed the ambiguity. In contrast, in the left+right-context condition, strong phonemic restoration effects left few consonants

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

unrestored. Half consonants were restored in complete breath groups, which included semantic, syntactic and prosodic information. The remaining consonants were restored at the word level, the prosodic-phrase level and the partial breath-group level. Missing the context of the previous sentence(s) was critical for some consonants, which remained unrestored. In both conditions, there was no effect of immediate vowel context.

It is known that natural speech production involves conveying a message via multiple paths of information and that in turn, perceptual recognition and sentence interpretation are achieved by making use of and integrating these sources of information. This implies a strong relationship between the production and perception levels. When words are clearly articulated and presented under ideal listening conditions, the context may simply accelerate word recognition (Morton & Long 1976). However, when words are poorly articulated and acoustic-phonetic information is impoverished and incomplete, much information from other sources is needed to restore the missing segments and phonemes. The results obtained here exemplify the latter tendency. Consonant restoration was found to be enhanced by lexical information and by the different sources of information contained in breath groups such as prosodic information, the preceding context (MacAllister 1988), and the following context (Grosjean 1985, Bard et al. 1988), which have proven to be involved in word recognition. The contribution of breath-group information was found to be fundamental here, suggesting that the breath group which is a production unit in conversational speech (Vaissière 1991) might also be a perception unit in this speech-style.

In the tasks described here, listeners had to restore assimilated and omitted consonants in gated intervals no greater than breath groups. Some of the information was thus removed. Moreover, although speech unfolds linearly in time, listeners were presented with the left-only context and with the left+right context. The acoustic and semantic properties of the stimuli and the specific cognitive demands of the task may have induced strategies that were different from those used in normal conditions of speech perception (Aaronson 1976). However, note above all that consonant restoration was usually achieved. One can assume that it would also have been quite successful with the previous and following sentence context. Interestingly, this finding is in agreement with what could be predicted in Lindblom's view (1990, 1996) that speech signals are perceptually adequate as long as they are rich enough to match the listener's access to signal-independent information, and that they need not be acoustically invariant but only sufficiently contrastive from the perceptual standpoint.

References

Aaronson, D. (1976). Performance theories for sentence coding: some qualitative observations. *Journal of Experimental Psychology: Human Perception and Performance* **2(1)**, 42–55.

Ali, L., Gallagher, T., Goldstein, J. & Daniloff, R. (1971). Perception of coarticulated nasality. *Journal of the Acoustical Society of America* **49**, 538–540.

Bard, E. G., Shillcock, R. C. & Altmann, G. T. M. (1988). The recognition of words after their acoustic offsets in spontaneous speech: effects of subsequent context. *Perception and Psychophysics* **44(5)**, 395–408.

Beddor, P. S., Krakow, R. A. & Goldstein, L. M. (1986). Perceptual constraints and phonological change: a study of nasal vowel height. *Phonology Yearbook* **3**, 197–217.

Benguerel, A. P. & Adelman, S. (1975). Coarticulation of lip rounding and its perception. In Cohen, A. & Nooteboom, S. E. (eds.), *Structure and Process in Speech Perception*, 283–293. Berlin: Springer and Verlag.

Byrd, D. (1996). Influences on articulatory timing in consonant sequences. *Journal of Phonetics* **24**, 209–244.

Cole, R. A. (1973). Listening for mispronunciations: a measure of what we hear during speech. *Perception and Psychophysics* **1**, 153–156.
Cole, R. A. & Jakimik, J. (1978). Perceptibility of phonetic features in fluent speech. *Journal of the Acoustical Society of America* **64(1)**, 44–56.
Duez, D. (1995). On spontaneous French speech: aspects of the reduction and contextual assimilation of voiced plosives. *Journal of Phonetics* **23**, 407–427.
Duez, D. (1998). Consonant sequences in spontaneous French speech. *Sound Patterns of Spontaneous Speech*, ESCA Workshop, La Baume-les-Aix, 63–68.
Foss, D. J. (1969). Decision process during sentence comprehension: effects of lexical item difficulty and position upon decision times. *Journal of Verbal Learning and Verbal Behavior* **8**, 457–462.
Fujimura, O. (1962). Analysis of nasal consonants. *Journal of the Acoustical Society of America* **34**, 1865–1875.
Grammont, M. (1914). *Traité Pratique de Prononciation Française*. Paris: Delagrave.
Grosjean, F. (1980). Spoken word recognition process and the gating paradigm. *Perception and Psychophysics* **28(4)**, 267–283.
Grosjean, F. (1985). The recognition of words after their acoustic offset: evidence and implications, *Perception and Psychophysics* **38(4)**, 299–310.
Lahiri, A. (1995). Undoing place assimilation. *Journal of the Acoustical Society of America* **97(3)**, 3333.
Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H & H theory. In Hardcastle, W. J. & Marchal, A. (eds.), *Speech Production and Speech Modelling*, NATO ASI Series, vol. 55, 403–439. Dordrecht: Kluwer.
Lindblom, B. (1996). Role of articulation in speech perception: clues from production. *Journal of the Acoustical Society of America* **99(3)**, 1683–1692.
Macallister, J. M. (1988). The use of context in auditory word recognition. *Perception and Psychophysics* **44(1)**, 94–97.
Mann, V. A. & Repp, B. H. (1980). Influence of vocalic context on the perception of the (ʃ)–(s) distinction. *Perception and Psychophysics* **28**, 213–228.
Miller, J. (1986). Rate-dependent processing in speech perception. In Ellis, L. (ed.), *Process in the Psychology of Language*, 119–157. Hillsdale, NJ: L. Erlbaum.
Morton, J. & Long, J. (1976). Effect of word transitional probability on phoneme identification. *Journal of Verbal Learning and Verbal Behavior* **15**, 43–51.
Ohala, J. J. (1986). Phonological evidence for top-down processing in speech perception. In Perkell, J. & Klatt, D. (eds.), *Invariance and Variability in Speech Processes*, 386–397. Hillsdale, NJ: Erlbaum.
Ohala, J. J & Feder, D. (1994). Listener's normalization of vowel quality is influenced by 'restored' consonantal context. *Phonetica* **51**, 111–118.
Ohala, J. J., Riordan, C. J. & Kawasaki, H. (1978). The influence of consonant environment upon identification of transitionless vowels. *Journal of the Acoustical Society of America* **64**, **S18**, J9.
Pollack, I. & Pickett, J. M. (1963). The intelligibility of excerpts from conversation. *Language and Speech* **6**, 165–171.
Rubenstein, H. & Pollack, I. (1963). Word predictability and intelligibility. *Journal of Verbal Learning and Verbal Behavior* **2**, 1947–1958.
Samuel, A. G. (1981). The role of bottom-up confirmation in the phonetic restoration illusion. *Journal of Experimental Psychology* **7(5)**, 1124–1131.
Samuel, A. G. (1987). Lexical uniqueness effects on phonemic restoration. *Journal of Memory and Language* **26**, 36–56.
Vaissière, J. (1991). Rhythm, accentuation and final lengthening in French. In Sundberg, J., Nord, L. & Carlson, R. (eds.), *Music, Language and Brain*, 108–120. Houndsmills: Macmillan.
Warren , R. M. (1970). Perceptual restoration of missing speech sounds. *Science* **167**, 393–395.
Warren, R. M. & Obusek, C. (1971). Speech perception and phonemic restorations. *Perception and Psychophysics* **9**, 358–363.

WHEELDON, L. & LAHIRI, A. (1997). Prosodic units in speech production. *Journal of Memory and Language* **37**, 356–381.
WIOLAND, F. (1991). *Prononcer les mots du français. Des sons et des rythmes*. Hachette: Paris.

Appendix

Pretest

/d/ reported as /n/
*des lieux de secon**d**e zone* 'second-rate places'

/d/ reported as /t/
*c'était le rideau **d**(e) fer* 'it was the iron-curtain'

/v/ reported as /f/
*il y a une chau**v**e-souris* 'there is a bat'

/s/ reported as omitted
*dans le s**t**yle roman* 'in the roman style'

Test

/d/ reported as /n/
*de gran**d**es manifestations de masse* 'huge mass demonstrations'
*y-a des points **d**(**e**) contact* 'there are contact points'
*c'est lié aux gran**d**es peurs* 'it is linked to the big fears'
*qu'ils se ren**d**en**t** compte* 'that they realize'
*dans les gran**d**es villes* 'in big cities'

/t/ reported as /n/
*ah la gen**t** féminine* 'ah the fair sex'
*de différen**t**es couleurs* 'with different colors'

/n/ reported as /m/
*toute u**n**e partie de la nuit* 'great part of the night'

/ʒ/ reported as /ʃ/
*si **j**(e) te l'avais dit* 'if I had told you'
*la **j**(e) pense que#* 'there I think that'
*et **j**(e) savais leur parler* 'and I did know how to talk to them'

/d/ reported as /t/
*non y-a pas **d**(e) souffrance* 'no, there is no pain'

/s/ reported as /z/
*là on **s**(e) demande* 'there one asks oneself'
*une espè**c**e de route* 'a kind of road'
*c'est la **s**vastika* 'it is the swastika'

/z/ reported as /s/
*la cri**s**e fait* 'the crisis makes'

/v/ reported as /f/
*qui ne doi**v**ent pas* 'who must not'

/p/ reported as /b/
*dans le grou**p**e des langues indo-européennes* 'in the Indo-European language group'

omitted /l/
*dans **l**(e) monde entier* 'in the whole world'
*Toulouse **l**a rose aussi* 'also Toulouse the pink city'

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press

*c'est qu'i**l**s ont essayé* 'but they did try'
*ses pu**l**sations cardiaques* 'one's heart beats'
*et i**l** nous a dit* 'and he told us'

omitted /ʁ/ or /ʀ/ (in the third example)
*c'est toujou**r**s systématique* 'it's always systematic'
*pou**r** les voyagistes* 'for travel organizers'
*la j(e) **r**epense aux mots* 'there I am again thinking of the words'

omitted /v/
*ou vi**v**re avec les tsiganes* 'or live with the gypsies'
*et j(e) **v**ais te dire* 'and I am going to tell you'
*qui peu**v**ent faire ça* 'who can do that'

omitted /z/
*de quin**z**e centimètres* 'fifteen centimeters long'

omitted /d/
*c'était quelque chose **d**'inoui* 'it was something unheard of'
*paquets **d**(e) biscuits* 'biscuit packets'

omitted /t/
*les plus hau**t**es décorations* 'the highest decorations'
*on n'inci**t**e pas les gens* 'People are not encouraged'
*c'est un peu s**t**ylisé* 'it is a little stylized'
*une boi**t**e de nuit sympa* 'A nice nightclub'

https://doi.org/10.1017/S0025100301001086 Published online by Cambridge University Press